Goss FR, Lai KH, Topaz M, Acker WW, Kowalski L, Plasek JM, Blumental KG, Seger DL, Slight SP, Fung KW, Chang FY, Bates DW, Zhou Li.

A value set for documenting adverse reactions in electronic health records.

***Journal of the American Medical Informatics Association* 2018**

DOI: https://doi.org/10.1093/jamia/ocx139

A Value Set for Documenting Adverse Reactions in Electronic Health Records

Foster R. Goss DO, MMSc[a], Kenneth H. Lai, MA[b] , Maxim Topaz, RN, MA, PhD[b,f], Warren W. Acker[b], Leigh Kowalski[b], Joseph M. Plasek, MS[b,c], Kimberly G. Blumenthal, MD, MSc[d,j], Diane L. Seger, RPh[e], Sarah P. Slight, MPharm, PhD[b,f,g], Kin Wah Fung, MD, MS, MA[h], Frank Y. Chang, MSE[h], David W. Bates, MD, MSc[b,f,j], Li Zhou, MD, PhD[b,i,j,*]

[a]University of Colorado, Department of Emergency Medicine, Aurora, CO, USA;
[b]Division of General Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA; USA;
[c]Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT, USA;
[d]Division of Rheumatology, Allergy and Immunology, and Medical Practice Evaluation Center, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA ;
[e]Clinical & Quality Analysis, Partners HealthCare System, Boston, MA, USA;
[f]Division of Pharmacy, School of Medicine, Pharmacy and Health, Durham University, Durham, UK; [g]Newcastle upon Tyne Hospitals NHS Foundation Trust, UK; [h]National Library of Medicine, Bethesda, MD, USA; [i]Clinical Informatics, Partners eCare, Partners HealthCare System, Boston, MA,USA; [j]Harvard Medical School, Boston, MA,USA;

Corresponding Author:
Foster Goss (foster.goss@ucdenver.edu)
Mail Stop B215, Leprino Building
12401 East 17th Ave
Aurora, Colorado 80045

Keywords: Allergic Reactions; Natural Language Processing; Allergy and Immunology; Electronic Health Records; Systematized Nomenclature of Medicine; Vocabulary, Controlled

Word Count: **4377**

ABSTRACT

Objective: To develop a comprehensive value set for encoding adverse reactions in the allergy module of an electronic health record (EHR).

Materials and Methods: We analyzed 2,471,004 adverse reactions stored in Partners Healthcare's enterprise-wide allergy repository (PEAR) of 2.7 million patients. Using the Medical Text Extraction, Reasoning, and Mapping System (MTERMS) system we processed both structured and free-text reaction entries and mapped these to SNOMED-CT. We then calculated the frequencies of reaction concepts, compared our value set to two external value sets, and then created an integrated value set. Lastly, we examined the presence of severe reactions and hypersensitivity reactions (HSRs) in our data.

Results: We identified 786 reaction concepts in PEAR. Most reported reactions were rash (14.0%), hives (8.2%), gastrointestinal irritation (5.5%), itching (3.2%) and anaphylaxis (2.5%). Mapping of Partners' value set to two other external sets identified 222 concepts that could be partially matched and 135 concepts that were missing. After incorporating missing and partial matches from external value sets, and removing duplicate concepts, our integrated value set included 1105 concepts. The presence of severe reactions was limited in both external sets. Hypersensitivity reactions represented roughly 20% of the reactions within our data.

Conclusion: We developed a value set for adverse reactions using a large dataset from one health system, enriched by reactions from two large external resources. This value set included severe and hypersensitivity reactions. We hope this work will improve reaction documentation and allergy-related clinical decision support.

OBJECTIVE, BACKGROUND, AND SIGNIFICANCE

Adverse reactions to foods, pharmaceuticals, and diagnostic products cause significant costs, morbidity and mortality in our healthcare system[1-3]. Adverse drug reactions (ADRs) have been reported to affect up to 10-20% of hospitalized patients and 25% of outpatients[4-6]. Some reactions, while rare, can be life threatening, for example toxic epidermal necrolysis (TEN), Stevens-Johnson syndrome (SJS), and immune hepatitis. In the United States, it has been estimated that nearly 1 in 300 hospitalized patients dies from an ADR every year[1]. Accurate documentation of a patient's adverse reactions to medications, products, or foods in the electronic health record (EHR) represents an important part of patient safety.

Allergy modules within EHRs provide a location for clinicians to document patients' adverse reactions. Although such modules use the term "allergy", many reactions entered in the allergy module are not immunologically mediated, though non-immunologically mediated reactions such as intolerances, toxicities, and idiosyncratic and pseudoallergic reactions can also be clinically important and often documented here. In fact, it is estimated that only about 5-10% of adverse reactions to drugs are allergic (immune-mediated)[7]. Some EHR systems provide coded options in the allergy module for clinicians to indicate a reaction as an allergy, intolerance, or contraindication, but studies have found that reaction type and classification are poorly understood by clinicians[8]. In addition, in most EHR systems, code sets for reactions are typically provided by third-party content vendors and may vary considerably based on the vendor and/or local management of terminologies. While useful, they are often limited in the number and granularity of reactions they contain, requiring users to enter a free-text reaction when the reaction they are looking for is missing. These free-text entries require information technologies such as natural language processing (NLP) to convert them to a coded form for subsequent automated processing.

Most prior efforts in standard terminologies have focused on capturing adverse events caused by drugs or devices for reporting [9 10] and pharmacovigilance purposes[11], rather than for clinical documentation in EHR systems. The World Health Organization Adverse Reaction Terminology (WHO-ART)[12] and Medical Dictionary for Regulatory Activities (MedDRA)[13] are used by pharmaceutical manufacturers for reporting adverse drug events (ADEs) to regulatory agencies. As they are mainly for reporting purposes, more clinical terminologies such as SNOMED CT have been suggested (e.g., by HL7 [14]) to encode reactions to an allergen or substance in EHRs. It is also worth mentioning that the International Classification of Diseases, Clinical Modification (ICD-CM) has been used to encode ADEs and hypersensitivity reactions as well with codes for rash, angioedema, Steven's Johnson syndrome and E-Codes to specify allergens [15-17].

Recently, an Adverse Clinical Reaction value set using SNOMED CT codes was provided by the National Library of Medicine (NLM) Value Set Authority Center (VSAC) with the support of the Federal Health Information Modeling and Standards (FHIMS) program. VSAC provides official versions of value sets used in clinical quality measures[18]. It also incorporates value sets for other use cases. The current published value set was comprised of reactions collected from Intermountain Healthcare, Kaiser Permanente, and the Veterans Administration's systems, based on reported frequencies. The value set spans 599 concepts stemming from multiple

Commented [PHS IS1]: We report ADRs to FDA (FAERS)

Commented [FRG2R1]: Should we add a comment that we report these?

hierarchies within SNOMED CT. To date, there are no published studies that perform external validation of this value set.

Therefore, we developed a new value set, including analysis of a large number of allergy entries from the Partners federated hospital/provider network EHR which have been used for some time, and then compared it to other value sets, to create a comprehensive value set. We use a NLP tool called the Medical Text Extraction, Reasoning, and Mapping System (MTERMS)[19 20] to process and map reaction terms to the Systematic Nomenclature of Medicine-Clinical Terms (SNOMED CT). Feasibility of this has been demonstrated in prior work on encoding food allergens[21], and extracting allergy information from clinical notes[20].

MATERIALS, METHODS, AND RESULTS

Our approach included of two phases (Figure 1). The first phase focused on processing reaction entries using NLP, creating a comprehensive reaction lexicon and mapping reactions to SNOMED-CT concepts. In the second phase, we calculated the frequencies of reaction concepts in our data set. We then compared our value set to FHIMS and a value set of reactions provided by the University of Nebraska Medical Center (UNMC), to create an integrated value set. Lastly, we examined the presence of severe reactions across each value set and the presence of hypersensitivity reactions (HSRs) in our own data. As our methods involved multiple steps and each step generated corresponding results, we present them in one section accordingly.

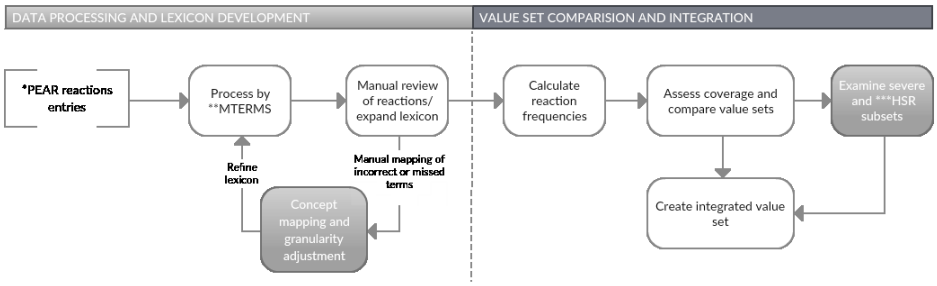

Figure 1. Overview of Methods
*PEAR: Partners Enterprise Allergy Repository
**MTERMS: Medical Text Extraction, Reasoning and Mapping System
***HSR: Hypersensitivity reactions

Definitions

Reactions are described at the entry, term, and concept level. We defined an "entry" as a word(s), phrase(s) or sentence(s) entered by a clinician into the reaction field of a patient's allergy record. A "term" was the portion of a reaction entry corresponding to the actual reaction, which is not normalized and may contain misspellings, acronyms, or syntactic variations. A "concept" was a collection of synonymous terms that represent a specific reaction. We defined a "value set" as a list of specific numerical values (codes) and human-readable names (concepts) derived from a standard terminology (i.e., in this study - SNOMED CT) within a specific clinical domain (i.e. in this case patients' adverse reactions) [22].

A "lexicon" is the vocabulary of our NLP system that is used to identify reaction terms in free-text entries. Our reaction lexicon contained additional terms collected from free-text that are not currently included in SNOMED CT. For example, in the reaction entry "he was switched from apap (acetaminophen) to asa (aspirin) due to elev (elevated) LFTs and cough", we extracted the terms "elev LFTs" and "cough". The term "elev LFTs" was mapped to the SNOMED CT concept ID 707724006 elevated liver enzymes level" (using its preferred term) along with other synonyms such as "LFT elevation".

We classified hypersensitivity reactions (HSRs) as *immediate* or *non-immediate*. Immediate HSRs have a time to onset less than one hour and are typically IgE-mediated, manifesting as urticaria, angioedema, rhinitis, conjunctivitis, bronchospasm, or anaphylaxis[23]. Non-immediate HSRs have a time to onset greater than one hour and are commonly T-cell mediated, manifesting with cutaneous symptoms including late-onset urticaria, maculopapular eruptions, fixed drug eruptions, vasculitis, toxic epidermal necrolysis, Stevens–Johnson syndrome, or drug reaction with eosinophilia and systemic symptoms syndrome[24 25].

Setting and Corpus

Our study used the Partners Enterprise-wide Allergy Repository (PEAR) which contains allergy information of all patients within the federated hospital/provider network, entered by clinicians in the EHR's allergy module [26]. As of October 26th, 2014, PEAR contained 3,949,996 active allergy entries for 2,730,250 unique patients, including drug, food, and environmental allergens, where 2,315,944 (58.6%) allergens had one or more reaction entries, accounting for 2,471,004 active allergy reaction entries. Among these reactions, 1,751,817 (70.9%) were coded entries (using 35 unique locally-defined codes, including "Unknown") and 719,187 (29.1%) were free-text entries. The average length of free-text entries was 12.1 [range 0-255] characters. Many free-text entries were long narratives containing other contextual information, such as "after a shot of PEN (penicillin), walked across the room and passed out, in childhood." To develop our lexicon, we utilized a subset of our corpus consisting of all entries with a frequency greater than 10 entries resulting in 539,610 (75.0% of total) free-text entries corresponding to 3,160 unique entries with an average length of 8.7 [range 0-68] characters.

To ensure the lexicon did not miss important but rare concepts, we randomly selected a subset of 500 reaction entries with a frequency of 10 or less (average length 22.1 [range 1-255] characters) for an internal evaluation. For external validation, we compared the adverse reaction value set we generated with two external value sets: 1) the FHIMS adverse reaction value set as described above; and 2) 604 unique adverse reaction concepts from the UNMC. This study was approved by the Partners and University of Colorado Multiple Institutional Review Boards (IRB).

Phase 1: Data Processing and Lexicon Development

Free-text reaction entries from PEAR were processed using MTERMS[19 27]. As described in prior work[19 21], MTERMS uses regular expression rules and a lexicon to: 1) process and tag allergens and reactions; 2) correct misspellings[28]; 3) handle contextual information (negations, temporality, exceptions, and abbreviations); and 4) map terms to concepts within standard terminologies. For this study, we mapped reaction terms in PEAR to the March 2016 release of

the U.S. Edition of SNOMED CT[29]. Data processing and lexicon development involved 4 steps as summarized in Figure 2 and described in detail below.

Step 1: Mapping Reaction Entries to SNOMED CT Descriptions
The first step focused on mapping PEAR reaction entries to SNOMED CT at the description (term) level. The terms found in this step formed the basis of our lexicon. MTERMS processed all 539,610 free-text reaction entries with a frequency greater than 10 (3,160 unique entries) and could map them to 757 SNOMED CT descriptions in the clinical finding hierarchy. The output was then manually reviewed by the study authors (FG, WA, KF) and pharmacy students who were in their 5th year of training. Thirty-one automatically mapped terms were not reactions (e.g., "male", "near") and therefore were removed from the lexicon, yielding 726 correctly mapped terms. We then manually mapped 1,109 terms that were unable to be automatically mapped (e.g., too poorly misspelled for the spell checker or use of abbreviations, e.g., "ITP" for "idiopathic thrombocytopenic purpura") to SNOMED CT. Finally, we identified four terms (itchy tongue, sores in throat, throat tingling and tingling in throat) that, although they represented reactions, we were unable to map to SNOMED CT, resulting in 1,835 terms in our lexicon.

Step 2: Mapping Reaction Terms to SNOMED CT Concepts
This step aimed to further expand the lexicon using the following three sub-steps. First we used MTERMS to map these 1,835 terms to their SNOMED CT concept IDs. For example, "rash" was mapped to two SNOMED concept IDs: 271807003 (eruption of skin (disorder)) and 112625008 (cutaneous eruption (morphologic abnormality)), because rash is listed as a synonym of these two concepts. Second, we included SNOMED CT synonyms for each identified concept. For example, for concept ID 271807003, we included the synonyms "eruption of skin", "eruption", "exanthema", etc. Third, we repeated the above two sub-steps until no more concept IDs or synonyms could be added. This resulted in 4,541 terms representing 834 concepts included in our lexicon.

Step 3: Refining Mapping Hierarchies
It is possible that a term was mapped to multiple concept IDs within different top-level hierarchies (or axes) of SNOMED CT. In the above example, "rash" was mapped to two concepts within two different hierarchies: disorder and morphologic abnormality; where the latter hierarchy is not related and therefore excluded from the lexicon. This enabled us to clean up the lexicon by excluding non-reaction concepts and unrelated hierarchies such as morphologic abnormality, qualifier value, and observable entity. At this stage, our lexicon comprised 4,300 terms representing 782 concepts split between the disorder (47.2%, n=369) and the findings concepts (52.8%, n=413).

Step 4: Assessing Lexicon Coverage by Processing Less Frequent Reaction Entries and Create Partners Value Set
To ensure we were not missing rare reactions with lower frequencies, we examined the coverage of our lexicon on a random subset of 500 entries from PEAR with a frequency less than or equal to 10. Using MTERMS, we processed this subset and calculated its coverage. Our lexicon performed well covering 96.6% of these concepts and supporting our belief that few new reactions were likely to be found in these longer narratives. MTERMS mapping performance was strong with a precision of 98.0%, recall of 95.3% and overall F-measure of 96.6%. With the

addition of the four concepts we were unable to map in Step 1, the total number of concepts in our lexicon rose to 786 concepts. These concepts formed the basis of our Partners Value Set (PVS).

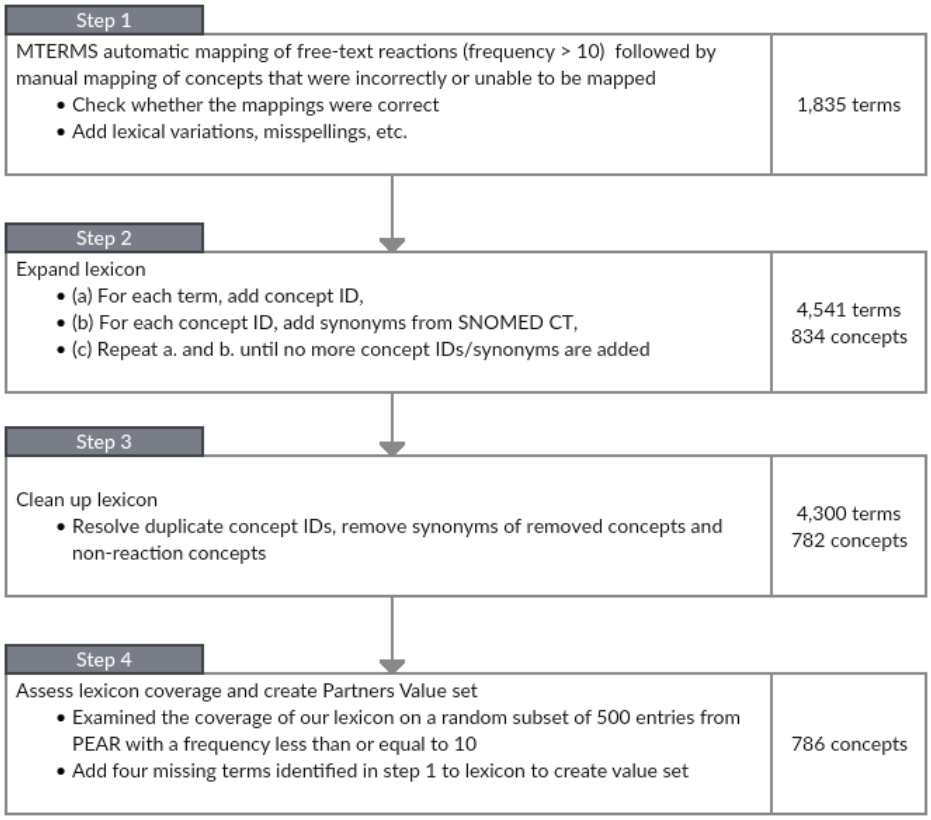

Figure 2. Adverse Reaction Lexicon Development

Phase 2: Frequencies of Reaction Concepts in PEAR

To assess the frequency of reaction concepts, we processed all reaction entries in PEAR (i.e. all free-text and structured entries) using MTERMS and our final lexicon. We identified in 2,584,112 terms (1,770,418 in the coded entries and 813,694 identified in the free-text entries), mapped these terms to their corresponding SNOMED CT concepts in the PVS, and calculated the frequencies of each concept in PEAR. We found the most frequently observed reactions were rash (eruption of skin) (13.96%, n=360,859), hives (weal) (8.25%, n=213,228), and gastrointestinal upset (irritation) (5.47%, n=141,389). A breakdown of the top 20 concepts between coded, free-text and overall frequency of entries is shown in Table 1 (full list in supplement A). Of interest, some of the most frequent concepts were only entered as free-text, including dizziness and palpitations.

Table 1: Top 20 Reaction Concepts

Reactions	SNOMED CT ConceptID	Total Frequency	Total %	Coded Frequency	Coded %	Free-text Frequency	Free-text %
Eruption of skin (i.e., Rash)	271807003	360,859	13.96	251,840	14.22	109,019	13.40
Weal (i.e., Hives)	247472004	213,228	8.25	165,108	9.33	48,120	5.91
Gastrointestinal irritation	95516005	141,389	5.47	113,323	6.40	28,066	3.45
Itching	418290006	81,462	3.15	67,707	3.82	13,755	1.69
Anaphylaxis	39579001	63,632	2.46	59,915	3.38	3,717	0.46
Nausea	422587007	45,300	1.75	20,035	1.13	25,265	3.10
Swelling	65124004	39,522	1.53	20,067	1.13	19,455	2.39
Cough	49727002	37,510	1.45	14,860	0.84	22,650	2.78
Vomiting	422400008	34,894	1.35	18,854	1.06	16,040	1.97
Angioedema	41291007	32,619	1.26	27,196	1.54	5,423	0.67
Altered mental status	419284004	29,709	1.15	27,788	1.57	1,921	0.24
Dyspnea	267036007	25,013	0.97	18,235	1.03	6,778	0.83
Muscle pain	68962001	23,568	0.91	8,539	0.48	15,029	1.85
Headache	25064002	22,534	0.87	8,084	0.46	14,450	1.78
Bronchospasm	4386001	18,521	0.72	17,934	1.01	587	0.07
Diarrhea	62315008	18,365	0.71	9,106	0.51	9,259	1.14
Wheezing	56018004	17,871	0.69	15,869	0.90	2,002	0.25
Sneezing	76067001	16,856	0.65	11,792	0.67	5,064	0.62
Dizziness	404640003	12,258	0.47	0	0.00	12,258	1.51
Palpitations	80313002	8,724	0.34	0	0.00	8,724	1.07
Others	-	403,556	15.62	36,664	2.07	366,892	45.09
Unknown	-	936,722	36.25	857,502	48.44	79,220	9.74
Total	**-**	**2,584,112**	**100**	**1,770,418**	**100**	**813,694**	**100**

Comparison with Existing Reaction Value Sets

Comparison among value sets was performed using methods that were similar to those we used in Zhou et al. [10], where we defined a match as exact, partial (broad or narrow) or missing. Broad matches in our study were defined when the terms in our lexicon were less specific than our target terminology (FHIMS, UNMC) (e.g., “Liver damage” -> “Cirrhosis of liver”). Narrow matches were defined when the term in our lexicon was more specific than the target terminology (e.g. “GI upset” -> “Gastrointestinal symptom”). Broad and Narrow matches were defined as involving only one reaction that could include a modifier (e.g. acute) or qualifier (e.g. severe). Where two or more reactions existed within a single concept, we classified this as pre-coordinated (e.g., nausea, vomiting and diarrhea). Both mapping and frequencies of reactions

within the source terminology were calculated. For example, in FHIMS, there were 92 partial matches that were classified as broad out of total of 599 representing 15.4% of concepts. Each concept is associated with a frequency in the source terminology. For example, the frequency of gastrointestinal symptom within UNMC was 0.113% (179 reaction entries/ 157,806 total reactions). These were summed across each value set and stratified by match type (Table 2).

When comparing FHIMS to PVS, at the *concept* level, we found 321 concepts within the two value sets having an exact concept ID match (Table 2). At the *term* level, there were 59 terms that were an exact match (e.g., bleeding from nose -> nose bleed) but with different concept IDs. In total, there were 380 (321+59) exact matches, representing 63.4% of the concepts in FHIMS. Among partial matches, there were 92 (15.4%) concepts in FHIMS that could be mapped to a broader concept in our PVS value set (e.g., respiratory crackles ->respiratory symptom) and 24 (4.2%) concepts that could be mapped to a narrower concept (e.g., serum cholesterol raised -> hypertriglyceridemia). There were 96 (16.0%) missing concepts where PVS did not have a match for the FHIMS concept. Pre-coordinated terms were small in number with 6 concepts that contained multiple reactions, representing 1.0% of concepts in FHIMS. By frequency of documentation, the number of concepts covered by our value set was 97.48% at the exact level and 99.36% including partial matches. Missing concepts represented 0.09%.

When comparing the UNMC reactions to PVS, at the *concept* level, there were 408 concepts that were an exact match by concept ID and 46 concepts that were an exact match at the *term* level. The total number of exact matches was 454, representing 75.2% of the concepts in UNMC. Among partial matches, there were 75 (12.4%) concepts in UNMC that could be could be mapped to a broader concept in PVS and 39 (5.1%) concepts that could be mapped to a narrower concept. There were 39 (6.5%) missing concepts and 5 (0.83%) pre-coordinated concepts. By frequency, the PVS value set covered 98.73% of concepts at the exact level and 99.78% including partial matches.

From the 135 (96+39) unique missing concepts, there were twelve (10 from FHIMS and 2 from UNMC) that upon review, we believe should be excluded from the integrated reaction value set. Reasons for possible exclusion are shown in Table 3.

Table 2: Concept coverage between PVS and FHIMS/UNMC

Match	**FHIMS**			**UNMC**			**Examples**
	# concepts	% concepts	% Frequency of concepts (source database)	# concepts	% concepts	% Frequency of concepts (source database)	
Exact Match	380	63.4	97.48	454	75.2	98.73	Heart irregular | irregular heart beat
Partial Broad*	92	15.4	1.55	75	12.4	0.72	Swelling of lower jaw region | face swelling
Partial Narrow**	24	4.17	0.32	31	5.1	0.33	Disease of liver | Liver damage

Missing	96	16.03	0.54	39	6.5	0.16	Mass, Elevated INR
Pre-coordinated	6	1.0	0.09	5	0.83	0.07	
Total (partial + exact)	**496**	**82.97**	**99.35**	**560**	**92.7**	**99.78**	

*Partial Broad: Broad matches in our study were defined when the terms in our lexicon were less specific than our target terminology (FHIMS, UNMC) (e.g., "Liver damage" -> "Cirrhosis of liver").
**Partial Narrow: Narrow matches were defined when the terms in our lexicon were more specific than the target terminology (e.g. "GI upset" -> "Gastrointestinal symptom")

Table 3: Reasons for possible exclusion

Concept	Source*	Frequency in Source	Reason for possible exclusion
Acute relapsing multiple sclerosis	FHIMS	0%	chronicity not necessary per se, use multiple sclerosis concept
Traffic accident on public road	FHIMS	0%	event, not a clinical finding or disorder
Malignant tumor of breast	FHIMS	0%	diagnosis, not a reaction
Fracture of femur	FHIMS	0%	Pathological fracture?, result of fall due to adverse reaction to medication?
Infectious disease	FHIMS	0%	too broad
Drug intolerance	FHIMS	0.0037%	classification of reaction, not a symptom of a reaction to an allergen
General health deterioration	FHIMS	0%	Too broad to be useful, rarely documented
Patients Condition Worsened	FHIMS	0%	Unclear what reaction worsened
Traumatic or non-traumatic injury	FHIMS	0%	Reaction type unclear - ? fall
Course of illness	FHIMS	0%	Concept is an attribute concept in SNOMED CT and should not be used to encode clinical information
Infection	UNMC	0.04%	Too broad, rarely documented
Treatment not tolerated	UNMC	0.001%	Reaction not specified, classified as intolerance

*Source database: referring to FHIMS or UNMC databases.

Comparison of Rare/Severe Reactions and Hypersensitivity Reactions (HSRs)

Next we estimated the presence of severe reactions and HSRs within our value set. Severe reactions, while low in frequency, can be the most critical for ensuring patient safety. Using a list of 15 severe reactions compiled by expert review (KB, FG) in PEAR, we examined the coverage of the FHIMS value set and UNMC corpus on these reactions. We found that FHIMS included 5 out of 15 of the rare and severe reaction concepts while UNMC included 6 out of 15 concepts (Table 4). Common among all were Stevens-Johnson syndrome and serum sickness, methemoglobinemia, and neuroleptic malignant syndrome. However, other important reactions, such as toxic epidermal necrolysis and drug reaction with eosinophilia and systemic symptoms (DRESS), while included in PEAR, but were not present in either the FHIMS or the UNMC data sets. Other severe reactions absent from FHIMS and UNMC included erythema nodosum, drug-induced hepatitis, fixed drug eruption, leukocytoclastic vasculitis, lichen planus, and aseptic meningitis.

HSRs were classified as likely immediate or non-immediate HSRs. Within our allergy repository, HSRs represented 19.2% of reactions (n=150 concepts). Common HSRs are shown below in table 5, divided among immediate and non-immediate.

Table 4: Presence of severe rare reactions among value sets

Reactions in PEAR	Freq in PEAR%	FHIMS*	UNMC**
Acute Interstitial Nephritis	0.0212%	Yes	No
Drug reaction with eosinophilia and systemic symptoms (DRESS)	0.0069%	No	No
Drug-induced hepatitis	0.0027%	No	No
Erythema multiforme	0.0195%	No	Yes
Erythema nodosum	0.0072%	No	No
Fixed drug eruption	0.0061%	No	No
Leukocytoclastic vasculitis	0.0027%	No	No
Lichen planus	0.0023%	No	No
Meningitis (aseptic)	0.0040%	No	No
Methemoglobinemia	0.0058%	Yes	Yes
Neuroleptic malignant syndrome	0.0070%	Yes	Yes
Pneumonitis	0.0073%	No	Yes
Serum sickness	0.0439%	Yes	Yes
Stevens-Johnson Syndrome (SJS)	0.0685%	Yes	Yes
Toxic Epidermal Necrolysis (TEN)	0.0104%	No	No

*FHIMS: showing the presence or absence of severe reactions in FHIMS value set.
**UNMC: showing the presence or absence of PEAR severe reactions in UNMC value set

Table 5: Hypersensitivity reactions

Immediate HSR*	Non-immediate HSR**
Anaphylaxis	Serum sickness
Hypotension	Drug rash eosinophilia and systemic symptoms syndrome
Angioedema	Erythema multiforme
Swelling	Acute interstitial nephritis
Rhinitis	Fixed drug eruption
Bronchospasm	Erythema nodosum
Asthma	Stevens-Johnson syndrome
Wheezing	Toxic epidermal necrolysis
Shortness of breath	Pneumonitis
Hives	Meningitis
Urticaria	Lichenoid drug eruptions
Itching	Leukocytoclastic vasculitis
Rash	

*Immediate HSR: Immediate HSRs (time to onset less than one hour) are typically IgE-mediated.
**Non-immediate HSR: Non-immediate HSRs (time to onset greater than one hour) are commonly T-cell mediated, but may be antibody or immune-complex-mediated.

Creation of an Integrated Value Set

Lastly, from the compared value sets, we then compiled an integrated value set that combines the 786 concepts in the PVS value set with 96 missing concepts from FHIMS and 39 missing UNMC concepts in addition to all the partial matches of each. In total, there were 1157 concepts. Each concept was manually reviewed (FG, KW, LK), duplicates removed (n=40) and excluded reactions above (n=12) removed, resulting in 1105 concepts. We recommend that these 1105

reaction concepts be included in an adverse reaction value set and IDs assigned to the concepts we could not find a concept ID in SNOMED CT.

DISCUSSION

We created a reaction value set by processing and analyzing reaction entries contained within an enterprise-wide large allergy repository using a semi-automated approach of NLP followed by manual review. Our value set covered between 63% and 75% of reactions by exact match in FHIMS and UNMC respectively. Based on frequency, these represented 97.5 to 98.7% of documented reactions. Partial matches represented 17-20% of reactions and the percentage of missing reactions ranged from 6-16% with a higher number of missing terms in FHIMS compared to UNMC. When integrated with existing value sets and curated to include missing concepts, partial matches, and remove ambiguous or duplicative concepts, the value set expanded from 786 unique concepts to 1105 reaction concepts. We believe this reaction value set captures the most frequently documented concepts, including severe and hypersensitivity reactions, collected from over two decades of documented reaction records, and that it could be a valuable tool for improving allergy documentation and allergy-related clinical decision support.

SNOMED CT for Encoding Reactions

We encountered several challenges during our evaluation including defining the appropriate hierarchy, reconciling ambiguous or duplicative concepts, and the use of the allergy section for documenting contraindications or religious preferences. With hierarchies, for example, a reaction to an allergen may include the *clinical finding* of irregular heart beat or the specific *disorder* of cardiac arrhythmia. While either might be correct, it would be desirable to default to one (e.g., clinical finding over the disorder) to avoid unintentional divergence in coding. Duplicative concepts were another challenge where there may be two concepts with a similar string but with a different concept ID. For example, the concept "Red Eye" was represented by two different concept IDs, one that referred to red eye as a "Finding of general observation of [an] eye" vs red eye as an "Ill-defined disorder of [the] eye". While debatable, the former was thought to be a better concept both in its ontological representation and by its frequency of documentation (825 entries vs. zero entries). While nearly every reaction term in our lexicon had a corresponding concept ID, some gaps did exist. A small number of reactions that were frequently documented were not present in SNOMED-CT, including 'itchy tongue', 'sores in throat', 'throat tingling', and 'tingling in throat'. Based on their frequency of occurrence, we would suggest adding these concepts, or a similar concept (e.g., oral itching) in SNOMED CT, with throat tingling and tingling in throat being represented as synonyms.

Value set Comparison

Overlap between PVS with FHIMS and UNMC reactions ranged between 82.9% and 92.7%. Manual review proved to be critical as concepts could be similar at the *term* level with different concept IDs, resulting in loss of content and coverage. One gap identified was in the coverage of severe reactions, which was notably lower in FHIMS and UNMC compared to PVS. These findings underscore the need for value sets to include severe reactions even though they may be low in frequency, highlighting the limitation of using frequency alone as a criterion for value set creation. With full representation of these rare/severe concepts in SNOMED CT, we would suggest including them in a reaction value set given their importance to patient safety and adverse event avoidance. Comparison of value sets allowed the identification of gaps present in each, their

missing or similar concepts, and the frequency of each concepts use. After adding partial and missing concepts, the integrated value set added approximately 500 concepts to FHIMS and UNMC, suggesting that the more than half of the content between value sets is at least similar at the *term* level. At the ontology level, future efforts will be needed to ensure consistency among concept IDs and perhaps default to one SNOMED-CT hierarchy to avoid divergence in coding.

Applications to Clinical Decision Support

With current override rates of 90% for inpatient and 77% of outpatient allergy alerts [30-33], redesigning CDS for allergy alerting could not be more important. Value sets for encoding reactions provide the necessary discrete data to re-align the type of CDS alert for a potential allergy or adverse reaction. This is particularly true for immune mediated responses which are associated with high risk. We found nearly 20% of reactions within our value set were HSRs (immediate and non-immediate). This estimate of HSR reactions may be conservative as reactions by themselves may not represent an HSR reaction but in combination with other reactions can be an HSR. For example, the GI symptoms nausea, vomiting, and diarrhea, in combination with shortness of breath may represent anaphylaxis (an HSR). Serum sickness may be described by the combination of fever, myalgias, arthralgias, and rash. Surveillance for reaction documentation for patterns can help clinicians identify rare but important reactions. Better understanding of the immunologic nature of a reaction (e.g. HSR) can help inform the type of alert to be presented to the clinician.

Documentation frequency of allergies can also be valuable for CDS. With known frequencies of reactions in PEAR and other value sets (FHIMS, UNMC), EHRs could conceivably populate a dynamic reaction quick pick list based on the most commonly associated reactions for a particular allergen. Using a data-driven approach to populate the reaction list could help limit the amount of time the clinician spends searching for a specific reaction, improve accuracy of documentation, and reduce inappropriate downstream alerting. The ideal EHR CDS module would know the most probable reactions for a given allergen, tier the allergy alerts based on: the reaction entered, severity, and prior results from allergy testing (skin tests, challenge results), and allow exceptions to account for medications previously tolerated or not cross-reactive. Contraindications would be handled using simple rule based advisories that could be triggered from any discrete data in the patient's record, be it religious preference, surgical history, or a problem on the patients problem list. Robust value sets for encoding patient reactions to allergens we believe are key to developing the infrastructure necessary to achieve a more intelligent advisory and alerting system for allergy and adverse reactions, limiting alert fatigue.

CONCLUSION

We processed and encoded reactions contained within a large allergy repository to validate and inform the maintenance of the Adverse Clinical Reaction value set maintained by the NLM and VSAC for interoperable use by EHRs. Our value set in addition to FHIMS will provide new insights into reaction and allergen associations, innovative CDS solutions for designing more intelligent allergy alerting, and improved allergy documentation.

Limitations

This work may be limited by the particular demographic of the population, which was primarily localized to the New England area. Allergens, particularly exposure to food and environmental allergens in this area may be different than those in other localities. While our reaction value set

contains the majority of reactions within PEAR, our lexicon was limited to those types of adverse reactions documented as free-text in the allergy module of the EHR. Other types of adverse reactions contained outside the EHR or in clinical notes may differ and enrich the reactions documented within in the allergy module.

CONTRIBUTIONS: All authors provided a substantial contribution to the conception and design of this work, its data analysis and interpretation, and helped draft and revise the manuscript. All authors are accountable for the integrity of this work.

ACKNOWLEDGEMENTS: Thank to James Campbell MD for sharing the UNMC reaction data. We also thank George Robinson, Shelly Spiro, and Robert McClure for their valuable advice on this study.

FUNDING: This study was funded by the Agency for HealthCare Research and Quality (AHRQ) grant R01HS022728.

COMPETING INTERESTS: None.

REFERENCES

1. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;**279**(15):1200-5
2. McDonnell PJ, Jacobs MR. Hospital admissions resulting from preventable adverse drug reactions. Ann Pharmacother 2002;**36**(9):1331-6
3. Patel DA, Holdford DA, Edwards E, Carroll NV. Estimating the economic burden of food-induced allergic reactions and anaphylaxis in the United States. J Allergy Clin Immunol 2011;**128**(1):110-15 e5 doi: 10.1016/j.jaci.2011.03.013[published Online First: Epub Date]|.
4. Gomes ER, Demoly P. Epidemiology of hypersensitivity drug reactions. Curr Opin Allergy Clin Immunol 2005;**5**(4):309-16
5. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. Br J Clin Pharmacol 2001;**52**(1):77-83
6. Gandhi TK, Weingart SN, Borus J, et al. Adverse drug events in ambulatory care. N Engl J Med 2003;**348**(16):1556-64 doi: 10.1056/NEJMsa020703[published Online First: Epub Date]|.
7. Frei CR, Attridge RT, Mortensen EM, et al. Guideline-concordant antibiotic use and survival among patients with community-acquired pneumonia admitted to the intensive care unit. Clinical therapeutics 2010;**32**(2):293-99
8. Blumenthal KG, Shenoy ES, Hurwitz S, Varughese CA, Hooper DC, Banerji A. Effect of a drug allergy educational program and antibiotic prescribing guideline on inpatient clinical providers' antibiotic prescribing knowledge. The Journal of Allergy and Clinical Immunology: In Practice 2014;**2**(4):407-13
9. Linder JA, Haas JS, Iyer A, et al. Secondary use of electronic health record data: spontaneous triggered adverse drug event reporting. Pharmacoepidemiology and drug safety 2010;**19**(12):1211-15

10. Goff DA, Bauer KA, Reed EE, Stevenson KB, Taylor JJ, West JE. Is the "low-hanging fruit" worth picking for antimicrobial stewardship programs? Clinical infectious diseases 2012:cis494
11. Avillach P, Coloma PM, Gini R, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. Journal of the American Medical Informatics Association 2013;**20**(1):184-92
12. HL7. FHIR. Secondary FHIR. https://www.hl7.org/fhir/overview.html.
13. International Conference on Harmonization. Secondary. http://www.ich.org/.
14. HL7. Secondary. https://www.hl7.org/fhir/allergyintolerance.html.
15. U.S. Department of Health and Human Services. International classification of diseases, 9th revision, clinical modifications (ICD-9-CM). 6th edition ed. Washington, DC: DHHS Publication No. (PHS) 96–1260., 1997.
16. Hougland P, Xu W, Pickard S, Masheter C, Williams SD. Performance of International Classification of Diseases, 9th Revision, Clinical Modification codes as an adverse drug event surveillance system. Medical care 2006;**44**(7):629-36
17. Hougland P, Nebeker J, Pickard S, et al. Using ICD-9-CM Codes in Hospital Claims Data to Detect Adverse Events in Patient Safety Surveillance. In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment). Rockville (MD), 2008.
18. National Library of Medicine. Value Set Authority Center. Secondary Value Set Authority Center 2015. https://vsac.nlm.nih.gov/.
19. Zhou L, Plasek JM, Mahoney LM, et al. Using Medical Text Extraction, Reasoning and Mapping System (MTERMS) to process medication information in outpatient clinical notes. AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium 2011;**2011**:1639-48
20. An Evaluation of a Natural Language Processing Tool for Identifying and Encoding Allergy Information in Emergency Department Clinical Notes. AMIA; 2014.
21. Plasek JM, Goss FR, Lai KH, et al. Food entries in a large allergy data repository. J Am Med Inform Assoc 2015 doi: 10.1093/jamia/ocv128[published Online First: Epub Date]|.
22. Graber CJ, Jones MM, Glassman PA, et al. Taking an Antibiotic Time-out: Utilization and Usability of a Self-Stewardship Time-out Program for Renewal of Vancomycin and Piperacillin-Tazobactam. Hospital Pharmacy 2015;**50**(11):1011-24
23. Sanchez E, Torres MJ, Mayorga C, et al. Adverse drug reactions with an immunological basis: from clinical practice to basic research. Allergy 2002;**57 Suppl 72**:41-4
24. Pichler WJ. Delayed drug hypersensitivity reactions. Ann Intern Med 2003;**139**(8):683-93
25. Posadas SJ, Pichler WJ. Delayed drug hypersensitivity reactions - new concepts. Clin Exp Allergy 2007;**37**(7):989-99 doi: 10.1111/j.1365-2222.2007.02742.x[published Online First: Epub Date]|.
26. Kuperman GJ, Marston E, Paterno M, et al. Creating an enterprise-wide allergy repository at Partners HealthCare System. AMIA Annu Symp Proc 2003:376-80
27. Goss FR, Zhou L, Plasek JM, et al. Evaluating standard terminologies for encoding allergy information. J Am Med Inform Assoc 2013;**20**(5):969-79 doi: 10.1136/amiajnl-2012-000816[published Online First: Epub Date]|.

28. Lai KH, Topaz M, Goss FR, Zhou L. Automated misspelling detection and correction in clinical free-text records. J Biomed Inform 2015;**55**:188-95 doi: 10.1016/j.jbi.2015.04.008[published Online First: Epub Date]|.
29. Terms) S-CSNoMC. Secondary. http://www.ihtsdo.org/snomed-ct.
30. Genco EK, Forster JE, Flaten H, et al. Clinically Inconsequential Alerts: The Characteristics of Opioid Drug Alerts and Their Utility in Preventing Adverse Drug Events in the Emergency Department. Ann Emerg Med 2016;**67**(2):240-48 e3 doi: 10.1016/j.annemergmed.2015.09.020[published Online First: Epub Date]|.
31. Topaz M, Seger DL, Lai K, et al. High Override Rate for Opioid Drug-allergy Interaction Alerts: Current Trends and Recommendations for Future. Stud Health Technol Inform 2015;**216**:242-6
32. Topaz M, Seger DL, Slight SP, et al. Rising drug allergy alert overrides in electronic health records: an observational retrospective study of a decade of experience. J Am Med Inform Assoc 2016;**23**(3):601-8 doi: 10.1093/jamia/ocv143[published Online First: Epub Date]|.
33. Slight SP, Beeler P, Seger DL, et al. Override Rates Of Drug Allergy Alerts In Inpatient And Outpatient Settings. BMJ Qual Saf 2016(0):1-9

Commented [LZ3]: Check this reference. It seems the first author is Hougland